REMARKS:

In the Office Action dated June 28, 2010, claims 1-16, in the above-identified U.S. patent application were rejected. Reconsideration of the rejections is respectfully requested in view of the above amendments and the following remarks. Claims 1-15 remain in this application, claim 16 has been canceled, claims 17-30 have been withdrawn and new claims 31-35 have been added to the application. New claims 31-35 are supported throughout the present application, for example, see original claim 1 and paragraphs [0017], [0019], [0020], [0144] and [0148].

Applicants thank the Examiner for the telephone interview on December 8, 2010. The amendments discussed regarding the 35 USC §101 rejection have been made in this response. The rejection under 35 USC §103(a) as unpatentable over Hoover was also discussed during the interview and the arguments discussed during the interview are included below.

The office action indicates that SEQ ID NOS are required throughout the specification. The specification was amended to include SEQ ID NOS in the Supplemental Preliminary Amendment Accompanying Filing of Sequence Listing filed on May 24, 2006.

The specification was objected to regarding section headings. The heading "Brief Description of the Drawings" has been added to the application. In addition, the hyperlinks have been deleted from the specification and the trademark GENBANK™ has been capitalized and marked as a trademark. In view of these amendments, applicants request that this rejection be withdrawn.

Claims 1-10 and 14-15 were rejected under 35 USC §101 as directed to nonstatutory subject matter. The claims have been amended to clarify that the method is tied to a suitably programmed computer and that the computer program is encoded on a non-transitory computer readable medium. In view of these amendments, applicants request that this rejection be withdrawn.

Claims 1-16 were rejected under 35 USC §103(a) as unpatentable over Hoover. In general, the Hoover paper describes a method of optimizing nucleotide sequences, wherein the sequence is broken into a number of contiguous sections and the codons are varied in each section. The result of a variation of codons in each iteration is accepted only as a whole. A sequence will always be accepted, if the quality function renders a better value (score) than the sequence accepted in a previous iteration. Additionally, a sequence having a worse value (score) will be randomly accepted with a certain probability, which will be varied during the iteration process. The new sequence is always adopted or rejected as a whole, never in part. In contrast to Hoover, in the presently claimed invention a test sequence is used as a starting sequence and m positions are designated where changes in the test sequence are to be made. These m positions are called optimization positions in the present claims and can change in each iteration. After a test sequence is generated, a new test sequence is created from the prior test sequence by exchanging one or more codons (maximum m) at the optimization positions for another codon. Then a further test sequence is created from the starting test sequence by making other replacements of one or more codons at the m optimization positions. This can involve replacing a codon by a different codon than before or replacing a codon at a different optimization position or both. This is repeated

to create further test sequences. This results in a large, but manageable number of test sequences that can be evaluated with a quality function. One of the test sequences is determined to be the optimum sequence with regard to this quality function. Then the result codons are determined. This means that from the m optimization positions, p positions are chosen as the positions of result codons. The codons that the optimum test sequence in this iteration has at these p positions will remain unchanged for all further iterations. In other words, once a result codon has been determined at a certain position, the codon at this position will remain unchanged during all subsequent iterations but any codons that are not result codons can be changed in the subsequent iterations. Thus, in the present invention, parts of the sequence are accepted, not just the whole sequence. This is clear from the claim language "where in each iteration step the test sequence comprises the appropriate result codon at the positions which correspond to positions of specified result codons in the optimized nucleotide sequence, and the optimization positions are different from positions of result codons" which indicates that the result codons are "accepted" and not changed in later iterations. The differences between Hoover and the presently claimed invention will be discussed in more detail below.

Hoover's purpose is the design of oligonucleotides for gene synthesis. Though optimization for expression in an organism is mentioned, it is not the primary purpose. The algorithm set out on p. 3 of Hoover proceeds as follows. First, the protein sequence is reverse-translated into a gene sequence by choosing the highest frequency codons in each case (p. 2, right column, last paragraph). The back-translated initial sequence is then divided into an odd number of contiguous sections, which are characterized by

near-equal melting temperatures. For each of these sections, a score is assigned, which takes into account the codon frequency, hairpin formation within each oligonucleotide, deviations from the desired melting temperature and size. For each section, an individual score is determined, which indicates the level of optimization. The score for the entire sequence is the sum of scores for all individual sections. For the optimization a stochastic method is adopted. The initial sequence is mutated by randomly swapping codons with others from the available pool of codons for the respective residue. The likelihood of being mutated can be higher in sections with high individual scores than in sections with a lower score. In each optimization step, the boundaries of the sections are also redefined. The result is a new (mutated) sequence, which corresponds to the entire protein. As is set out in more detail in Hoover on p. 3, middle of the left column and top of right column, the score is calculated for each section of the new sequence and a new overall score is calculated by adding the individual scores of the individual sections. If the overall score of the new sequence improves over the old sequence, the new sequence is kept. If it does not improve, the new sequence is adopted with a certain likelihood (determined by a Boltzmann distribution) and with the remaining likelihood the old sequence is maintained. As discussed above the sequence is always adopted or rejected as a whole, never in part. Thus, Hoover's process is a stochastic process. Within one iteration, the selection of the codon to be changed is random and the choice of the replacement codon is also random. The search space for finding the optimum overall sequence, which covers essentially all possible combinations of codons representing a given amino acid sequence, is not investigated systematically. The chance of finding the optimal

sequence (having the best theoretically possible score) with regard to local characteristics is very low. This is due to the fact that intermediate results of one iteration are not fixed or saved in Hoover. It is possible that in one iteration a residue is changed from codon A to codon B and in a subsequent iteration changed back from codon B to codon A.

Comparing the algorithm of Hoover to the method of the present claim 1, the following differences are apparent.

- In the present invention, optimization positions are specified, i.e. in every iteration the sequence is mutated only at the specified optimization positions, which are a subset of the codons to be optimized, not at random codons and at randomly selected positions of the entire sequence as in Hoover.
- According to Hoover the result of an iteration is not necessarily the sequence with the best score, but can be a sequence with a worse score than a previous sequence with a likelihood that is determined by a Boltzmann distribution. In contrast, the codons chosen as result codons, i.e. those codons that are not varied in subsequent iterations and thus form part of the final optimized sequence are chosen from the optimal test sequence found in an iteration.
- According to Hoover's algorithm, in every iteration the sequence is either maintained as a whole or rejected as a whole, but never maintained in part as in the present invention.
- Hoover discloses a global optimization approach while the present invention is directed to a more localized optimization approach. From a biological perspective the sequence properties which influence its technical usefulness (e.g. maximal expression

of the encoded protein in a host organism) are less global properties (like the overall GC content) but much more local features (e.g. the GC content within a window of 40 base pairs or the presence/absence of a certain sequence motif, which may influence the expression). The present invention addresses this issue by looking at only a subpart of the sequence within one iteration and finding the best possible solution for this part with a non-stochastic approach, e.g. by doing an exhaustive search for the optimal subsequence. After the optimal subsequence is determined it is not changed in later iterations.

- The determination of result codons which are not changed in later iterations is not suggested or disclosed by Hoover. In the present invention, once the optimal occupation of the m optimization positions is found, a number (p) of these optimal codons are fixed as result codons and are not changed in the course of further iterations. This means that in subsequent iterations of the algorithm the m optimization positions will be different from previously established result codons. In other words the presently claimed process employs kind of a "memory" to preserve the optimal solution for a subpart of the whole sequence throughout further iterations. In contrast, the only information Hoover's approach saves unalterably throughout the multiple iterations is the quality score of the sequence generated in the last iteration immediately before the current iteration. Hoover's method randomly varies codons in different iterations.

Applicants respectfully contend that the cited prior art does not suggest or disclose a localized approach as in the present claims nor was it shown that such approaches were common knowledge to one skilled in the art at the priority date of the present application. However, even if this had been the case, there is no suggestion

regarding the specific approach of claim 1, specifically, an approach comprising the steps of selecting m optimization positions, p result codons and maintaining result codons in further iterations. This is in fact supported by Hoover. On p. 3, left column, last paragraph, Hoover points out that the number of possible sequences (2^{100} for a protein with 100 residues assuming only 2 codons per amino acid residue, which is in the order of magnitude of 10^{31}) is such that there is no way of calculating all possible variations. This is the reason why Hoover used a stochastic method. In Hoover's method all variables, which would ultimately be codons, can be varied at any stage of the optimization process. Referring to the numbering of the present invention method steps in the office action (page 10), applicants respectfully point out that there is no disclosure of step 5 in Hoover, namely that (1) in each iteration a number (p) of codon sites are determined at which the codon occupation will not be varied in subsequent iterations over the optimum test sequence in the current iteration and (2) the codon occupation is determined by the codon occupation of the test sequence in this iteration having the best score.

Page 11 of the office action states that "Hoover et al. stress that during each optimization step, the number of mutations as well as changes in sizes diminish as the optimization process progresses". Applicants assume that this statement refers to the paragraph bridging the left and right columns of page 3 of Hoover. This text indicates that the likelihood of mutation is higher in sections of a high individual score than in those having a lower score (a lower score means a better sequence in the context of Hoover). This is different from step 5 of present claim 1. First of all, this text relates to process steps within one iteration, whereas step 5 is a step performed at the end of

each iteration. This text passage teaches away from the present invention, because in the present invention, in one iteration there are fixed optimization positions and the number of these optimization positions is not changed within a single iteration. According to the Hoover paper both the positions where codons are varied and the number of positions at which codons are varied changes in a stochastic manner within a single iteration. In addition, this does not change the requirement that in Hoover's method, at the end of an iteration the optimized sequence will be rejected or accepted as a whole, whereas according to the present invention part of the sequence can be accepted and the rest of the sequence can be further varied. Thus, Hoover is directed only to a global approach and does not suggest or disclose the more localized approach used in the present invention.

The presently claimed invention has several advantages over the prior art. From a computation stand point, the approach of the present invention is less complicated and it in fact permits the calculation of the exact solution to the problem of optimizing the codons at the m optimization positions in each iteration step. Because you limit the variation of codons to a number m of sites in the present invention, the number of all possible combinations of codons in an iteration step remains manageable and it is possible to evaluate any possible codon combination on the m sites. Thus, local correlations or interdependencies that are between on the scale of m codons will be considered exactly and thus the exact local optimum on the m sites can be determined. Nevertheless, since the quality function is evaluated not only over the m optimization positions, but over the entire test sequence, long-range correlations or

interdependencies, such as the GC content, can be taken into account and will influence the optimization result.

A further disadvantage of the statistical method suggested by Hoover is that due to its statistical nature, it does not provide a unique result. In other words, if you run the same optimization procedure twice with the same parameters, the result may be different. Thus, the result achieved by Hoover's method is not necessarily the optimum and one skilled in the art would not know whether they have actually achieved the optimum codon sequence or not.

Another drawback of the statistical method of Hoover is that the codons are changed at random and only a small number of all possible sequences are actually considered. Since the optimization is done on the basis of a global function, desired local correlations such as the occurrence of a succession of certain specific nucleotides on a small scale are unlikely to be found. This is because the likelihood that an exact succession of given nucleotides is found in a random process is small, as the probabilities multiply. Moreover, even if in the course of the optimization process a sequence is generated that contains the desired succession of nucleotides, it could be discarded with a certain probability in subsequent iterations, even if it has a better score than a new sequence. Regarding the specific algorithm disclosed in Hoover, one skilled in the art would assume that it would be impossible to generate sequences having a predetermined short succession of specific nucleotides. From a biological viewpoint, this is important with regard to sequence motifs, for example, short sequences of specific nucleotides are desired in the sequence in order to generate restriction sites. Hoover's algorithm is unlikely to produce such specific successions of nucleotides or

sequence motifs. Dinucleotides are also unlikely to be produced by Hoover's method (for further reference see WO 20061015789 A2 corresponding to CA 2 575 480 Al, which relates to the targeted insertion or removal of CpG dinucleotides). In contrast to Hoover, according to the present invention, if the number m is greater or equal to the length of the desired sequence or motif, one can calculate all possible codon combinations on said m optimization positions and the desired sequence motif, e.g. the desired dinucleotide, can be found with certainty.

Another biological issue is the interdependency between GC content and codon frequency. In certain organisms, such as dicotyledons, codon frequency and GC content are conflicting targets, i.e. optimizing the GC content will deteriorate the optimization of codons in terms of highest frequency and vice versa. Thus, for such organisms one has to make a compromise between the codon frequency and the GC content. In the present invention, at every step, the codon choice can be made in such a way that the best combination of codon frequency and GC content can be found. Thus, in each iteration, the test sequence is optimized in terms of GC content and codon frequency. Therefore the sequence converges towards an optimum combination of GC content and codon frequency in each iteration. In contrast to the present invention, in Hoover's method, an improvement in terms of GC content and codon frequency, can only occur by the above mentioned random swapping of one codon for another in order to generate a new test sequence. Since the use of a low frequency codon in order to improve the GC content at one location may be compensated by the use of one or more codons with a higher frequency at other locations and it is not known

at which positions the necessary changes are to be made, it is unlikely that a process in which codons are swapped at randomly chosen sites will lead to the very sequence that embodies the best compromise between GC content and codon frequency. In other words, given the huge number of possible sequences (2^{100}), the Hoover algorithm only checks a very small number of sequences, which are chosen at random, hence it is unlikely that a specific sequence embodying the best compromise will be found. In contrast to Hoover, in the present invention, in every iteration the result codons that provide the optimum in terms of GC content and codon frequencies can be chosen and fixed so that they are not changed in subsequent iterations. Thus, in the Hoover method, the likelihood that an optimum combination of GC content and codon frequency is produced, is much lower than in the present invention and it is much more likely that you will end up with a sequence that is either unsatisfactory in term of codon frequency or in terms of GC content (or both).

Applicants respectfully point out that the present invention permits the optimization of a sequence on a small scale, e.g. on the scale of a few nucleotides, in such manner that desired sequence motifs or other local features, such as CpG dinucleotides, restriction sites, can be included in the resulting sequence while simultaneously optimizing properties that are related to the sequence in its entirety, such as GC content. Though optimization results will vary depending on how the parameters m and p and the score are chosen, one can generally say that using the method of the presently claimed invention, it is possible to provide the exact optimum on the scale of m positions while taking into account long-range interdependencies. In

view of the differences between the presently claimed invention and Hoover, applicants request that this rejection be withdrawn.

Enclosed with this response are the claims which were granted in Australia and Japan in applications corresponding to the present application.

Applicants respectfully submit that all of claims 1-15 and 31-35 are now in condition for allowance. If it is believed that the application is not in condition for allowance, it is respectfully requested that the undersigned attorney be contacted at the telephone number below.

In the event this paper is not considered to be timely filed, the Applicant respectfully petitions for an appropriate extension of time. Any fee for such an extension together with any additional fees that may be due with respect to this paper, may be charged to Counsel's Deposit Account No. 02-2135.

Respectfully submitted,

By /Monica Chin Kitts/
Monica Chin Kitts
Attorney for Applicant
Registration No. 36,105
ROTHWELL, FIGG, ERNST & MANBECK
1425 K. Street, Suite 800
Washington, D.C. 20005
Telephone: (202) 783-6040

MCK/cb